CELLULAR AUTOMATA BASED DNA CRYPTOGRAPHY ALGORITHM

G. Shanmuga Sundaram
Dept. of Information Technology
Sri Manakula Vinayagar Engineering College
Madagadipet, Puducherry
shanmugasundaram@smvec.ac.in

S.Pavithra , A.Arthi, B.Madu Bala, S.Mahalakshmi
Dept. of Information Technology
Sri Manakula Vinayagar Engineering College
Madagadipet, Puducherry
s.pavithra2693@gmail.com, arthi2694@gmail.com ,
madu54am@gmail.com, masha.lak25@gmail.com

Abstract— **DNA cryptography is a new area of research in cryptographic field. In this DNA component is used as the information carrier. Most of the encryption techniques based on the cellular automata have limitations. To overcome this lacuna, we propose a novel DNA cryptography algorithm with cellular automata to achieve randomness, parallelism, uniformity, reversibility and stable. Finally a comparison is made with the research works based on the cryptographic attack parameter.**

Index Terms—Cellular automata, DNA, Thymine, Uracil

I. INTRODUCTION

DNA cryptography is a new promising field in cryptography which emerged with the progress of DNA computing. It can be used in two ways to encrypt and decrypt, first is to encrypt the data as DNA sequence and second is converting that DNA sequence into DNA structure which can be embedded into any organism like amoeba and it can be sent instead of transferring encrypted data [1]. But our work concentrates on to use only DNA sequence for the process. It will provide high level security. Cellular automata can be applied in several areas in cryptographic field; it has much property to support for effective cryptographic algorithm [12]. Along with the DNA cryptography, cellular automata can be combined to form an effective algorithm.

II. BACKGROUND WORK

The DNA cryptography is a very promising direction in cryptography research. DNA can be used in cryptography for storing and transmitting the information, as well as for computation. Here we try to utilize the power of DNA to hide message [1]. The massive parallelism and extraordinary information density inherent in this molecule are exploited for cryptographic purposes, and several DNA based algorithms is proposed for encryption, authentication and so on. The current main difficulties of DNA cryptography are the absence of theoretical basis, the high tech lab requirements and computation limitations.

DNA Cryptography

The DNA cryptography is a very promising direction in cryptography research. DNA can be used in cryptography for storing and transmitting the information, as well as for computation. Here we try to utilize the power of DNA to hide message [1]. The massive parallelism and extraordinary information density inherent in this molecule are exploited for cryptographic purposes, and several DNA based algorithms is proposed for encryption, authentication and so on. The current main difficulties of DNA cryptography are the absence of theoretical basis, the high tech lab requirements and computation limitations[2].

Various DNA Encryption Techniques

Cryptography and data security are critical aspects of conventional computing. Private key encryption method which is secure even in presence of a computationally unbounded is the One-Time-Pads (OTP) cipher. OTP is secure in theory. But practically, key distribution and generation are critical problems for usage of OTP cryptosystem. DNA has high memory density, it resolve the large cipher key producing and saving issues [3].

Gehani [4] proposed two encryption methods. First method is to convert the DNA plain code sequence cell to DNA cryptograph sequence according to mapping graph defined; it is called as mapping substitute. Other method is exclusive or method, it uses biological molecular techniques to carry through exclusive OR operation of DNA plain code and key sequence. These two methods used in OTP encryption mechanism.

According to Andre Leier there are two different cryptographic approaches based on DNA binary strands. The one approach hides information in DNA strands and the second approach is molecular checksum [4].

The message is encoded and hidden in DNA dummies which can be restored by PCR amplification[3], followed by sequencing operations as a one-way function have a method for key distribution based on public-key system using DNA proposed by kazoo Tanaka.

A recombinant DNA based cryptosystem [4] was presented by Nini Rao. Encryption schemes based on DNA computing mostly utilize its handling step, structure or biochemistry condition.

978-1-4799-6480-2/15/$31.00 ©2015 IEEE

Since there are limited works exist in DNA cryptography field and it is still in maturing state. From these research contribution it is clearly found that there exist a wide opportunity in the use of DNA in the cryptography mechanism.

Cellular Automata

"A system made up of many discrete cells, each of which may be in one of a finite number of states. A cell or automaton may change state only at fixed, regular intervals, and only in accordance with fixed rules that depend on cells own values and the values of neighbors within certain proximity. "

Formula Definition

CA = one- or two-dimensional grid of identical automata cells**.** Each cell processes information and proceeds in its actions depending on its neighbors**.** Each cell (automaton) A defined by

- Set of States S = $\{S_1, S_2, S_3, S4\ldots, S_N\}$
- Transition Rules T
- Therefore A ~ (S,T,R) (R: neighboring automata)
- T: $(S_t, I_t) \rightarrow S_{t+1}$

Cellular Automaton

A one-dimensional cellular automaton (CA) consists of two things: a row of "cells" and a set of "rules". Each of the cells can be in one of several "states". The number of possible states depends on the automaton. Think of the states as colors. In a two-state automaton, each of the cells can be either black or white. Over time, the cells can change from state to state. The cellular automaton's rules determine how the states change. It works like this: When the time comes for the cells to change state, each cell looks around and gathers information on its neighbors' states. Based on its own state, its neighbors' states, and the rules of the CA, the cell decides what its new state should be. All the cells change state at the same time [14].

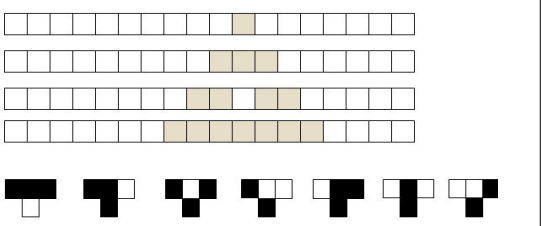

Fig. 1. Cell Definition

3 Black = White
2 Black = Black
1 Black = Black
3 White = White

There are 256 automata rules from rule 0 to rule 255 each rule has its own pattern and characteristics.

Cellular Automata in Cryptography

This section discuss about the existing cryptographic techniques based on cellular automata.

The study of DES algorithm with cellular automata (DES) was proposed by Rama R et al.,[7] which exhibit the capability of producing random and complex behaviour. CA is a parallel processing machine and it is context dependent behaviour of all values at time.

From their research work it is clearly stated that usage of CA rules are simple in nature for the operation of DES and AES algorithms. The proposed CA- based key generation uses the elementary cellular automata rule 30 which possess more randomness. Key size used is minimal in this approach

Pratibha Sharma et al., proposed the Text security using 2D cellular automata rules. It provides two levels of security. First level comprises of hiding message to be sent behind the image using password and second level comprises of encryption using 2D CA rules. It ensures more security to transmit message. Two level securities are due to chip level in automata machine, the process of encryption and decryption is in high speed [8].

Triple Stage DNA cryptography using sequential machine was proposed by Rupalisoni et al., to secure data there are traditional techniques in cryptography such as RSA, DES, and AES are used. This work represents a new approach to secure the data which is based on DNA cryptography & Automata theory. DNA cryptography technique was proposed by L.M.Adleman in 1994 for solving Hamiltonian path problem. Automata theory was initiated in 20th century. These two techniques are combined to form a powerful algorithm[9].

From the various research contributions, it is clearly states that DNA based cryptosystem can be designed using cellular automata concept. But the existing research contribution are seems to be limited.

III. PROPOSED WORK

Our objective of this work is to develop a highly secured system and support the needs of the user. The study of both DNA cryptography and celluar automata fields inferred that their characteristics which will lead to provide effective cryptographic algorithms. So the algorithm designed which satisfy the above characteristics of the field. This algorithm comprises the technique of both DNA cryptography and cellular automata. Here the plain text is encrypted and decrypted using these techniques. Initially the plain text is taken from the user either in file or direct input. The input file is intially convert data in the file into DNA sequence until the end of file. Then convert the DNA sequence into binary file and divide the binary sequences into parts and perform xor operation on the binary parts with the help of key on the n parts after xor operation. Convert the Resultant value of xor operation i.e binary sequence into DNA sequence and send to automata machine here with the help of automata machine the DNA sequence is convert into DNA code words which is in the form of text and number and in unrecognized form. Thus

along with the result output embedded the key value and its generation used as cipher text to the receiver. The algorithm flow is shown in the flow diagram shown in fig.2.

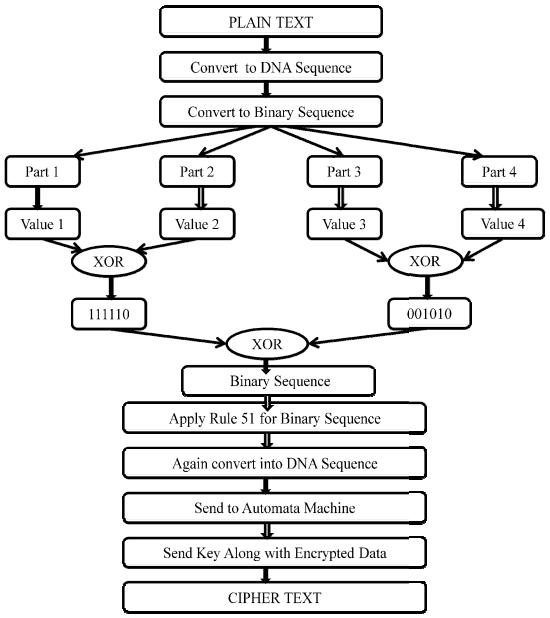

Fig.2. Flow Diagram of Proposed Algorithm

Proposed Algorithm

Input: *F[file of any type], A[Automata for Plain text to AGTC data generation], c1[code book1], c2[code book2], c3[code book3], c4[code book4]*

Algorithm : Encryption

Step 1: *The File F is converted into DNA sequence:*

F⟶ F[DNA(U)] ; // AGTC sequence file.

Step 2: *In the File F[DNA] convert the Uracil into Thymine acid*

F[DNA(U)] ⟶ F[DNA(T)];

Step 3: *The DNA sequence is converted into binary sequence, The DNA output file of*

F[DNA(T)]⟶Binary[F[DNA(T)]];
// Binary file

Step 4: *Split the Binary Sequence into equivalent parts:*

*Binary[F[DNA(T)]]⟶ Binary[**F1**[DNA(T)]]*
*⟶ Binary[**F2**[DNA(T)]]*
*⟶ Binary[**F3**[DNA(T)]]*
*⟶ Binary[**F4**[DNA(T)]]*

Step 5: *Perform XOR operation with help of keys K1, K2, K3, K4 for all Binary Sequences F1, F2, F3, F4. // key generated and used.*

*X1=**(**Binary[**F1**[DNA(T)]]**)** (XOR)K1*
*X2=**(**Binary[**F2**[DNA(T)]]**)** (XOR)K2*
*X3=**(**Binary[**F3**[DNA(T)]]**)** (XOR) K3*
*X4=**(**Binary[**F4**[DNA(T)]]**)** (XOR) K4*

Step 6: *Combine the all output of the binary sequence into a single sequence.*

X= X1+X2+X3+X4;

Step 7: *Apply the rule 51 for the resultant binary sequence X.*

Now the state of x in time t is X(t).

RULE 51 (00110011) ⟶ s(t+1) = NOT(s_currentCell(t))

Next State of X(t) is X(t+1).

X(t+1) = Rule 51(X(t));

After applying rule, resultant will be binary sequence. // Initial state t is transformed into t+1 next state.

Step 8: *Convert the Binary sequence into DNA sequence*

X(t+1)⟶ X[DNA(t+1)];
S= X[DNA(t+1)]

Step 9: *In the Sequence S convert the Thymine(T) into Uracil(U) acid.*

S⟶ S(T);

Step 10:*Send the S(T)⟶ Automata Machine A[S(T)]; // DNA code word is generated*

Step 11: *Repeat the above steps until reach the end of the file;*

Step 12: *The Encrypted Data is the output of the automata machine*

A[S(T)]⟶ Encrypted Sequence;

Step 13: *Append the Key in the output file at the end*

Encrypted File[F]⟶ Encrypted Sequence + key;

This output file Contains the DNA code sequence.

Output: *File F with the AGTC coded DNA sequence*

End Algorithm.

Stages in Algorithm

Here as shown above in fig 3.1, the process is divided in to three major stages.

1) Initial stage
2) Middle Stage
3) Final Stage

Initial stage: In this stage the plain text which is taken as input from the user is converted into DNA sequence by the codebook1 and codebook 2 given below.

Code Book 1

Table 1: Conversion of Plain text to DNA Sequence

Sl.No	Plain Text	DNA Sequence	Sl.No	Plain Text	DNA Sequence
1	A	UUU	14	N	GUC
2	B	UUC	15	O	GUA
3	C	UUA	16	P	GUG
4	D	UUG	17	Q	UCU
5	E	CUU	18	R	UCC
6	F	CUC	19	S	UCA
7	G	CUA	20	T	UCG
8	H	CUG	21	U	CCU
9	I	AUU	22	V	CCC
10	J	AUC	23	W	CCA
11	K	AUA	24	X	CCG
12	L	AUG	25	Y	ACU
13	M	GUU	26	Z	ACC

Code Book 2

Table 2: Conversion of Numbers and Special Characters to DNA Sequence

S. No	Number & Symbols	DNA Sequence	S. No	Number & Symbols	DNA Sequence
1	1	ACA	21	{	GAA
2	2	ACG	22	}	GAG
3	3	GCU	23	[	UGU
4	4	GCC	24	]	UGC
5	5	GCA	25	\	UGA
6	6	GCG	26	|	UGG
7	7	UAU	27	“	CGU
8	8	UAC	28	‘	CGC
9	9	UAA	29	:	CGA
10	0	UAG	30	;	CGG
11	!	CAU	31	?	AGU
12	@	CAC	32	/	AGC
13	#	CAA	33	>	AGA
14	$	CAG	34	.	AGG
15	%	AAU	35	<	GGU
16	^	AAC	36	,	GGC
17	&	AAA	37	_	GGA
18	*	AAG	38	-	GGG
19	(	GAU	39	+	UU1
20	)	GAC	40	=	CC1

Then convert DNA sequence into binary sequence using the codebook 3.

Code Book 3

Table 3: Conversion DNA Sequence to Binary Sequence

Sl.No	Digital Code word	DNA Code word
1	00	A
2	01	G
3	10	T
4	11	C

Middle Stage: Divide the Binary sequence into equal parts and perform XOR operation with help of key and in resultant binary XOR value apply rule 51 of cellular automata given figure 3.[15]:

Features of rule 51

1. Uniformity
2. Reversibility
3. Stable

Pattern of Rule 51

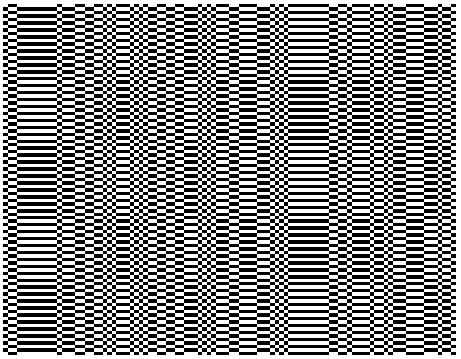

Fig. 3. Rule 51

Final Stage: This stage contains two phases to encrypt data.

(1) First phase: convert the Binary sequence into AGTC data with the help of the table3.

(2) Second phase: Send the AGTC data to Automata machine in the fig 4.

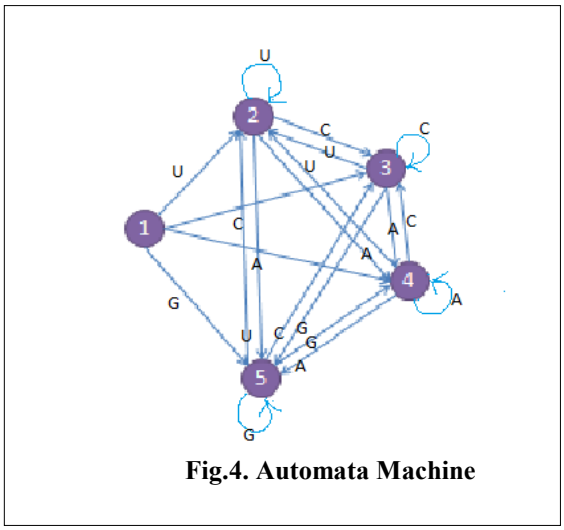

Fig.4. Automata Machine

With the help of the code book 4 and automata machine convert the DNA sequence into DNA code word.

Code Book 4

Table 4: DNA code to automata code word

S.no	DNA code	Code Word	S.no	DNA code	Code Word
1	UUU	F1	34	UAC	Y2
2	UUC	F2	35	UAA	B1
3	UUA	L1	36	UAG	B2
4	UUG	L2	37	CAU	H1
5	CUU	L3	38	CAC	H2
6	CUC	L4	39	CAA	Q1
7	CUA	L5	40	CAG	Q2
8	CUG	L6	41	AAU	N1
9	AUU	I1	42	AAC	N2
10	AUC	I2	43	AAA	K1
11	AUA	I3	44	AAG	K2
12	AUG	M1	45	GAU	D1
13	GUU	V1	46	GAC	D2
14	GUC	V2	47	GAA	E1
15	GUA	V3	48	GAG	E2
16	GUG	V4	49	UGU	C1
17	UCU	S1	50	UGC	C2
18	UCC	S2	51	UGA	B1
19	UCA	S3	52	UGG	W1
20	UCG	S4	53	CGU	R1
21	CCU	P1	54	CGC	R2
22	CCC	P2	55	CGA	R3
23	CCA	P3	56	CGG	R4
24	CCG	P4	57	AGU	S5
25	ACU	T1	58	AGC	S6
26	ACC	T2	59	AGA	R5
27	ACA	T3	60	AGG	R6
28	ACG	T4	61	GGU	G1
29	GCU	A1	62	GGC	G2
30	GCC	A2	63	GGA	G3
31	GCA	A3	64	GGG	G4
32	GCG	A4	65	UU1	X1
33	UAU	Y1	66	CC1	X2

Finally append the key with the DNA code word and send to receiver, that is encrypted text data or file.

IV. THEORETICAL ANALYSIS OF PROPOSED ALGORITHM

Here we are concluding the result on the basis of the nature of the proposed mechanism and the various types of the security attacks. The study shows that the proposed mechanism is much secured and has less threat from various known attacks.

Table 5: Theoretical Analysis of cellular automata based cryptography algorithm on various attacks

S.no	Attack Type	Defence	Remarks
1	Cipher text attack	Strong	Strong defence due to triple encryption
2	known Plain Text Attack	Strong	Strong defence due to Automata replacement of original text
3	Chosen Plain Text Attack	Strong	Very difficult because DNA implementation
4	Chosen Cipher Text Attack	Strong	Very difficult due to Double DNA encryption
5	Adaptive Chosen Plain Text	Strong	Very difficult due to DNA encryption
6	Adaptive Chosen cipher Text	Strong	Very difficult due to AGTC code to amino acid code

The various attacks on plain text and cipher text are theoretically described and discussed in table 5.

V. CONCLUSION

In this work a novel DNA Cryptographic algorithm with cellular automata to perform encryption and decryption is proposed. This algorithm is secured at 4 levels of XOR operations, conversion in DNA, DNA is converted into binary based sequence, and it is subjected to rule 51 of cellular automata. The theoretical analysis of attack resistant illustrates the strength of the algorithm. In future, the more cellular rule can be applied to develop an effective algorithm.

References

[1] U.Noorul Hussain, T.Chithralekha, "Review of DNA cryptology", CiiT International Journal of Networking and Communication Engineering, vol 3, No 13, October 2011, pg No 843-849.

[2] Grasha Jacob, A. Murugan, "DNA based Cryptography: An Overview and Analysis", ISSN: 2222-4254, Int. J. Emerg. Sci., 3(1), 36-42, March 2013.

[3] H. Z. Hsu and R. C. T. Lee, "DNA Based Encryption Methods", The 23rd workshop on combinatorial mathematics and computation theory.

[4] Guangzhao Cui, Cuiling Li, Haobin Li, Xiaoguang Li "DNA Computing and Its Application to Information Security Field", 2009 fifth international conference on natural computation.

[5] S.Jeevidha, Dr.M.S.SaleemBasha, Dr.P.Dhavachelvan "Analysis on DNA based Cryptography to Secure Data

Transmission" International Journal of Computer Applications (0975 – 8887) Volume 29– No.8, September 2011.

[6] Sambhu Prasad Panda, MadhusmitaSahu, Umesh Prasad Rout, Surendra Kumar Nanda "Encryption and Decryption algorithm using two dimensional cellular automata rules in Cryptography", International Journal of Communication Network & Security, Volume-1, Issue-1, 2011.

[7] Rama R, Bala Suyambu J, Andrew Arokiaraj, Shanmugam Saravanan, "A Study of DES Algorithm with cellular automata", *International Journal of Innovative Management, Information & Production* ISME International ©2011 ISSN 2185-5439 Volume 3, Number 1, March 2012 PP. 10-16 .

[8] Pratibha Sharma, NiranjanLal, ManojDiwakar "Text Security using 2D Cellular Automata Rules", Conference on Advances in Communication and Control Systems 2013 (CAC2S 2013).

[9] RupaliSoni*, GopalPrajapati, Arif Khan, Deepak Kulhare, "Triple Stage DNA Cryptography Using Sequential Machine", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 8, August 2013 ISSN: 2277 128X.

[10] M.Venugopal, Dr. E.G.Rajan, Dr. Sharma, "Security measures for Smart Devices through Cryptography Using Cellular Automaton", M.Venugopal et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (1) , 2013, 129 – 133.

[11] PradiptaMaji and Chandrama Shaw, NiloyGanguly, Biplab K. Sikdar and P. Pal Chaudhuri, "Theory and Application of Cellular Automata For Pattern Classification", FundamentaInformaticae 58 (2003) 321–354 321,IOS Press.

[12] Stephen Wolfram, "Random sequence generation by cellular automata". Advances in Applied Mathematics, 7:123{169}, 1986.

[13] Cellular automaton - Wikipedia, the free encyclopedia

url:http://en.wikipedia.org/wiki/Cellular_automaton retrieved on july 15, 2013

[14] Cellular Automata rules lexicon - User DLLs -

url:http://psoup.math.wisc.edu/mcell/rullex life.html retrieved on Aug 26, 2013.

[15] Cellular Automaton -- from Wolfram MathWorld -

url:http://mathworld.wolfram.com/UniversalCellularAutomaton.html retrieved on aug 31, 2013.

[16] Cellular Automata The 256 Rules (Stanford Encyclopedia of Philosophy)- url:http://www.roguebasin.com/index.php retrieved on september 16, 2013.